

ESTIMATING SOIL PARTICLE-SIZE DISTRIBUTION FOR SICILIAN SOILS

Vincenzo Bagarello, Vito Ferro, Giuseppe Giordano

1. Introduction

The particle-size distribution (PSD) of the soil is a fundamental soil physical property and it is routinely measured to classify soils. Many investigations have shown that the PSD is also the basis for estimating soil hydraulic properties such as the water retention curve [Arya 1981; Fredlund 1994; Fredlund 2002; Haverkamp 1986; Saxton 1986]. A conventional particle-size analysis involves measuring the mass fractions of clay, silt and sand and using these fractions to classify the soil by a textural diagram. A more complete description can be obtained by using a theoretical PSD model.

The PSD of a soil has been assumed to be approximately lognormal [Campbell 1985; Shirazi 1984] but many other models have been proposed [Buchan 1989; Buchan 1993; Fredlund 2000; Jaky 1944; Nemes 1999; Shiozawa 1991]. The PSD models differ by their fitting ability, and the performances of a particular model may vary with the soil textural characteristics [Hwang 2004; Hwang 2002]. The selected model may have a significant impact on the estimated soil-particle percentage at a given particle-size limit [Nemes 1999], but only a few comparative studies of PSD models have been carried out in soil science [Hwang, 2004]. Buchan [1993] verified the adaptability of the lognormal distribution to 71 measured PSDs collected in New Zealand. Hwang [2002] and Hwang [2004], using a data set of 1387 Korean soils, concluded that the Fredlund [2000] (F) model had the best performance with the majority of investigated soils. A practically important implication of this result is that a single PSD model can be used in large areas including soils with different textural characteristics. However, the conclusions by the Korean investigation have to be tested in other parts of the world. Another model needing further testing is the one by

Haverkamp [1986] (HP), that was derived from the van Genuchten [1980] function representing water retention characteristics of soil. A particular form of the HP model was used in the recently proposed BEST (Beerkan Estimation of Soil Transfer parameters) method to facilitate the hydraulic characterization of unsaturated soils [Lassabatère 2006].

Fitting a theoretical PSD model to the measured one has been carried out using databases with different number of sampled particle diameters [Fredlund 2000; Haverkamp 1986; Hwang 2002, Skaggs 2001]. Minasny [2007] suggested that at least five measurements of particle size fractions are needed to characterize PSD properly, but no detailed information is available on the relationship between the number of measured data points and the estimated model parameters.

The objectives of this study were to i) evaluate the ability of the HP and F models to fit 243 measured PSDs from a wide range of soil textures in Sicily, and ii) test the effect of the number of measured particle diameters on the fitting of the theoretical PSD.

2. Materials and methods

The Sicilian data-base contained 243 soil samples uniformly distributed throughout the island (Fig.1), with a sampling density approximately equal to one sample per 100 km².

For each sampling point, the PSD was measured using conventional methods following H₂O₂ pretreatment to eliminate organic matter and clay deflocculation by sodium metaphosphate and mechanical agitation. Fine size fractions were determined by the hydrometer method, whereas the coarse fractions were obtained by mechanical sieving. In particular, sieving analysis was carried out using six sieves with mesh sizes of 2000, 860, 425, 250, 106 and 75 µm. Eight fine fraction data points were obtained by the hydrometer method, measuring the suspension density at 2, 5, 15, 30, 60, 180, 1440 and 2880 min. The minimum diameter measured by the hydrometer method varied with the sample from 0.000821 to 0.000997 mm.

Paper received 27.01.2009; accepted 08.05.2009

Dipartimento di Ingegneria e Tecnologie Agro-Forestali, Università di Palermo, Viale delle Scienze, 90128 Palermo.

Particle-size fraction data were classified according to the USDA standards [Gee 1986]. The Sicilian soils data-base represented a wide range of soil textures. Table 1 lists the number of measured PSDs for each USDA textural class whereas Fig. 2 shows the textural composition of the Sicilian soils considered in this investigation.

For representing the measured PSD, the HP and F theoretical distributions were used. In particular, the HP model [Haverkamp 1986] is a functional relationship of the van Genuchten [1980] type:

$$P(d) = \left[1 + \left(\frac{d_g}{d} \right)^n \right]^{-\frac{1-g}{n}} \quad (1)$$

where $P(d)$ is the fraction by weight of soil particles having a diameter less than or equal to d (L), d_g (L) is the scale parameter of the distribution and n is a shape parameter always assuming values greater than 1. In particular, an increase of d_g shifts the curve towards higher diameter values whereas the distribution tends to become horizontal as n decreases ($P(d)$ always equal to one for $n = 1$).



Fig. 1 - Location of the soil sampling points in Sicily.

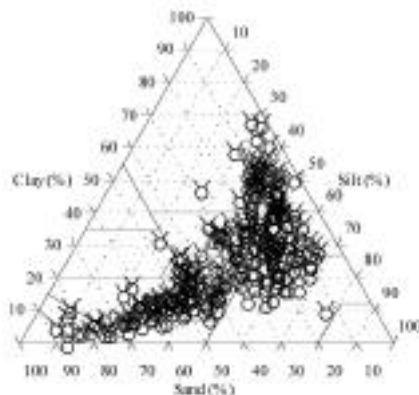


Fig. 2 - Textural composition of the soil data set according to the USDA classification system.

The F model presented by Fredlund [1994] and Fredlund [2000; 2002] in the unimodal form is the following:

$$P(d) = \frac{1}{\left[\ln \left(\exp(1) + \left(\frac{\alpha}{d} \right)^{n_F} \right) \right]^m} \left\{ 1 - \frac{\ln \left(1 + \frac{d_f}{d} \right)}{\ln \left(1 + \frac{d_f}{d_m} \right)} \right\} \quad (2)$$

in which d_m (L) is the diameter of the minimum allowable particle size, d_f (L) is a parameter related to the amount of fines in a soil, α (L) is a parameter designating the inflection point on the curve, n_F is a parameter related to the steepest slope on the curve (uniformity of the PSD), and m is a parameter related to the shape of the curve as it approaches the fines region [Fredlund et al., 2000]. According to the parametric analysis of eq.(2) developed by Fredlund [2000], it follows that: i) if α increases, while the other parameters are held constant, the PSD of eq.(2) moves towards larger particle sizes; ii) the n_F parameter influences the slope of the PSD; iii) the curve moves towards the smallest particle sizes when m decreases; and iv) large variations of the d_f parameter produce a minimal influence on the curve. According to this last result, Fredlund [2000] suggested to use a constant d_f value, equal to 0.001 mm. In the investigations by Hwang [2002] and Hwang [2004], d_f was considered to be a fitting parameter but d_m was set equal to 0.001 mm [Hwang 2002] or 0.0001 mm [Hwang 2004].

For each soil sample, the two-parameter HP model (eq. 1) was fitted to the measured PSD using a non-linear estimation technique. For the F model, taking literature indications into account, α , n_F , m and d_f were used as fitting parameters and $d_m = 0.0001$ mm, which is lower than the minimum practically measurable diameter, was assumed (F4 scenario). The F model was also tested fixing d_f at 0.001 mm, as suggested by Fredlund [2000], and d_m at 0.0001 mm (F3 scenario). This last fitting procedure reduced eq.(2) to a three-parameter unimodal PSD model. The estimation of the three parameters of the F model (F3 scenario) was also carried out using eight measurements, i.e. six measurements obtained by sieving and the first (at 2 min) and last (at 2880 min) hydrometer measurements (F3/8 scenario).

Therefore, the selected theoretical PSDs were characterized by two (HP model), three (F model, F3 and F3/8 scenarios) or four (F model, F4 scenario) parameters, with the aim to define the optimum distribution as the best model fitting the data with a low number of fitting parameters. This choice can be justified taking into account that the fitting performance generally improves as the number of fitted parameters increases even if the risk of overparameterization increases, too [Hwang 2002].

The model's performances were evaluated by the following indices:

$$MXE = \max_{i=1}^N [abs(P_i(d) - F_i(d))] \quad (3)$$

$$ME = \sum_{i=1}^N \frac{P_i(d) - F_i(d)}{N} \quad (4)$$

$$RMSE = \sqrt{\sum_{i=1}^N \frac{[P_i(d) - F_i(d)]^2}{N}} \quad (5)$$

where N is the number of the measured pairs (diameter, d - frequency by weight, $F_i(d)$) and $P_i(d)$ is the corresponding theoretical probability calculated by the selected model. The maximum error, MXE , is a local indicator of the goodness of the estimation provided by a given theoretical PSD model. The mean error, ME , reveals the presence of a possible bias (i.e., a systematic overestimation or underestimation of the fitted values). Finally, the root mean square error, $RMSE$, quantifies the scatter of the data points around the PSD model.

3. Results and discussion

The MXE , ME and $RMSE$ values obtained with the F model and the F4 scenario (α , n_F , m , and d_f ; fitting parameters; $d_m = 0.0001$ mm) were compared with the

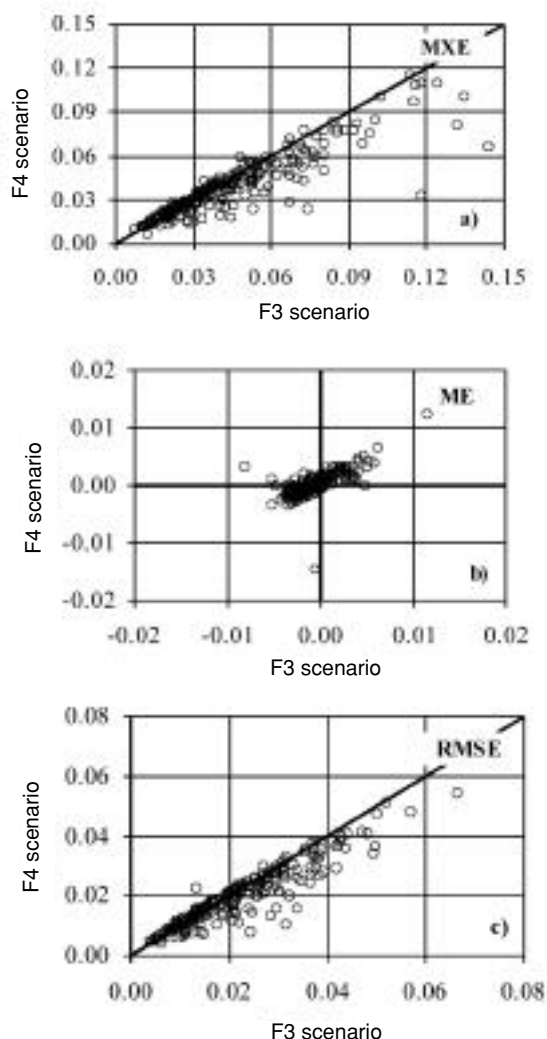


Fig. 3 - Comparison between the MXE , ME and $RMSE$ values obtained for the F model with two different estimating procedures of the d_f parameter (F4 and F3 scenarios).

corresponding values obtained by the same model and the F3 scenario (α , n_F , and m : fitting parameters; $d_f = 0.001$ mm; $d_m = 0.0001$ mm) (Fig. 3).

For the MXE and $RMSE$ indices, the data pairs were generally close to the 1:1 line. For the ME index, a systematic underestimation or overestimation was not associated with the use of a particular scenario. Therefore, using the F3 scenario instead of the F4 one made the fitting only slightly worse and confirmed that the d_f parameter did not influence substantially the performance of the F model [Fredlund 2000]. Therefore, only the simpler F3 scenario was considered for subsequent analysis.

Note that with $d_f = 0.001$ mm, as suggested by Fredlund et al. [2000], and $d_m = 0.0001$ mm [Hwang 2004], the term included into braces of eq. (2) is very close to 1 (in particular, it ranges from 0.999 to 1 for $0.00064 \leq d \leq 2$ mm). This result suggested that the theoretical PSD may be written in the following simplified form:

$$P(d) = \frac{1}{\left[\ln \left(\exp(1) + \left(\frac{\alpha}{d} \right)^{n_F} \right) \right]^m} \quad (6)$$

Fig. 4 compares the values of MXE , ME and $RMSE$ calculated with the F model (F3 scenario) and the corresponding values obtained for the HP model. Generally, the two-parameter HP model yielded higher MXE , $abs(ME)$ and $RMSE$ results than the three-parameter F model, showing a better performance of this last model. This result was expected only in part, because other investigations have shown that increasing the number of parameters cannot always guarantee improved performances [Hwang 2002]. The F model yielded lower mean values of MXE and $RMSE$ and a mean ME result closer to zero than the HP model for each soil textural class (Table 1). Therefore, the F model showed the best fitting ability, as compared with the HP one, for all textural classes included in this investigation (Fig.5).

Establishing in absolute terms the goodness of the performance of a selected PSD model is difficult because guidelines to classify the MXE , ME and $RMSE$ results have not been developed. For a given PSD, a mean deviation between the measured and the predicted percentages by not more than 3% (i.e., $RMSE \leq 0.03$) could be considered to be practically negligible.

According to this approximate criterion, the HP model yielded, on average, satisfactory results for the clay and silty-clay soil textural classes, whereas the F model performed well in the clay, silty-clay, silty-clay-loam, silt-loam, clay-loam, loamy-sand, and loam soils (Table 1).

Therefore, both models performed better in fine-textured soils than in coarse-textured soils, but only the F model yielded satisfactory results for the majority of the investigated soils, as also concluded by

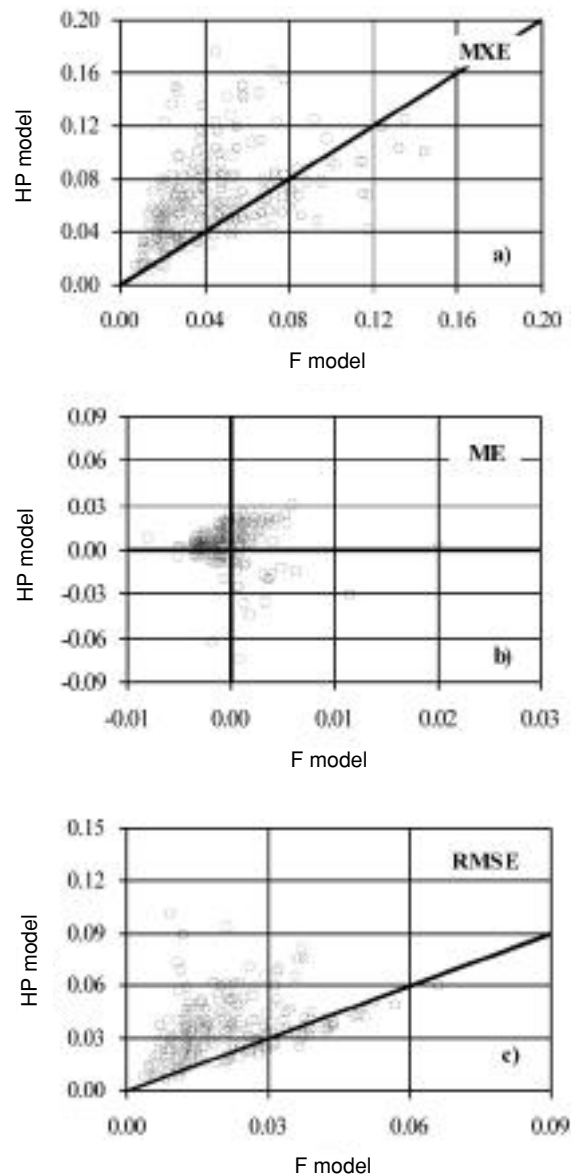


Fig. 4 - Comparison between the *MXE*, *ME* and *RMSE* values obtained for the F (F3 scenario) and HP models.

Hwang [2002] for the Korean soils.

Table 1 also lists the mean values of *MXE*, *ME* and *RMSE* calculated for each soil textural class using the F model with the F3/8 scenario. As expected, using eight measured data pairs determined a fitting of the theoretical distribution to the complete measured distribution worse than the F3 scenario (Fig. 5). A direct comparison between the F3 and F3/8 scenarios in terms of *ME* and *RMSE* results was not possible because a different number of data points was used in the two cases ($N = 14$ and $N = 8$, respectively).

However, the eight data points considered in the F3/8 scenario were a sub-set of the complete data set used in the F3 scenario. If a higher *MXE* result is obtained with the F3/8 scenario, then the conclusion must be that using a limited number of data pairs to fit eq.(2) determines a maximum discrepancy between measured and predicted frequencies higher than the one obtained by the complete data set. In this investigation, the F3/8 scenario yielded, on average, larger *MXE* results than the F3 one for eight soil textural classes out of ten (Table 1). Moreover, the number of experimental data points used in the fitting procedure affected the ranking of model's performances by soil textural class. For example, the worst performance in terms of *MXE* and *RMSE* results was detected for the loam class when $N = 8$ data points were used and the sand one with $N = 14$ (Table 1).

Finally, according to the $RMSE < 0.03$ criterion, satisfactory results were obtained for four textural classes (silt-loam, silty-clay-loam, clay, silty-clay) when a limited experimental information was used in the fitting procedure, which was a different result than the one obtained with $N = 14$ data points. Therefore, this investigation suggested that the number of data points used in the fitting procedure may have an appreciable impact on the detected model's performances. The number of measured particle-size fractions is a factor that should be taken into account when different investigations have to be compared.

Soil textural class	Sample size	F model (F3 scenario)			HP model			F model (F3/8 scenario)		
		<i>MXE</i>	<i>ME</i>	<i>RMSE</i>	<i>MXE</i>	<i>ME</i>	<i>RMSE</i>	<i>MXE</i>	<i>ME</i>	<i>RMSE</i>
Loam	50	0.0594	-0.0013	0.0287	0.0666	0.0020	0.0343	0.0732	-0.0162	0.0382
Sandy loam	38	0.0645	-0.0004	0.0315	0.0830	-0.0069	0.0385	0.0667	-0.0065	0.0350
Silt loam	30	0.0341	0.0001	0.0164	0.0645	0.0137	0.0370	0.0620	0.0003	0.0282
Silty clay loam	29	0.0285	0.0009	0.0144	0.0564	0.0118	0.0305	0.0543	-0.0004	0.0245
Clay loam	29	0.0427	-0.0015	0.0217	0.0634	0.0061	0.0358	0.0658	-0.0175	0.0346
Clay	28	0.0230	-0.0007	0.0121	0.0405	0.0057	0.0215	0.0536	-0.0108	0.0260
Silty clay	25	0.0241	0.0003	0.0124	0.0445	0.0078	0.0238	0.0642	-0.0078	0.0297
Loamy sand	10	0.0629	0.0016	0.0254	0.1164	-0.0312	0.0623	0.0606	0.0024	0.0307
Sand	2	0.0794	0.0029	0.0328	0.1135	-0.0191	0.0520	0.0684	0.0011	0.0362
Sandy clay loam	2	0.0638	-0.0001	0.0304	0.0989	-0.0089	0.0508	0.0646	-0.0144	0.0357

TABLE 1 - Mean values of the *MXE*, *ME* and *RMSE* statistics calculated by using the models by Fredlund et al. [2000] (F3 and F3/8 scenarios) and Haverkamp and Parlange [1986] for each USDA soil textural class.

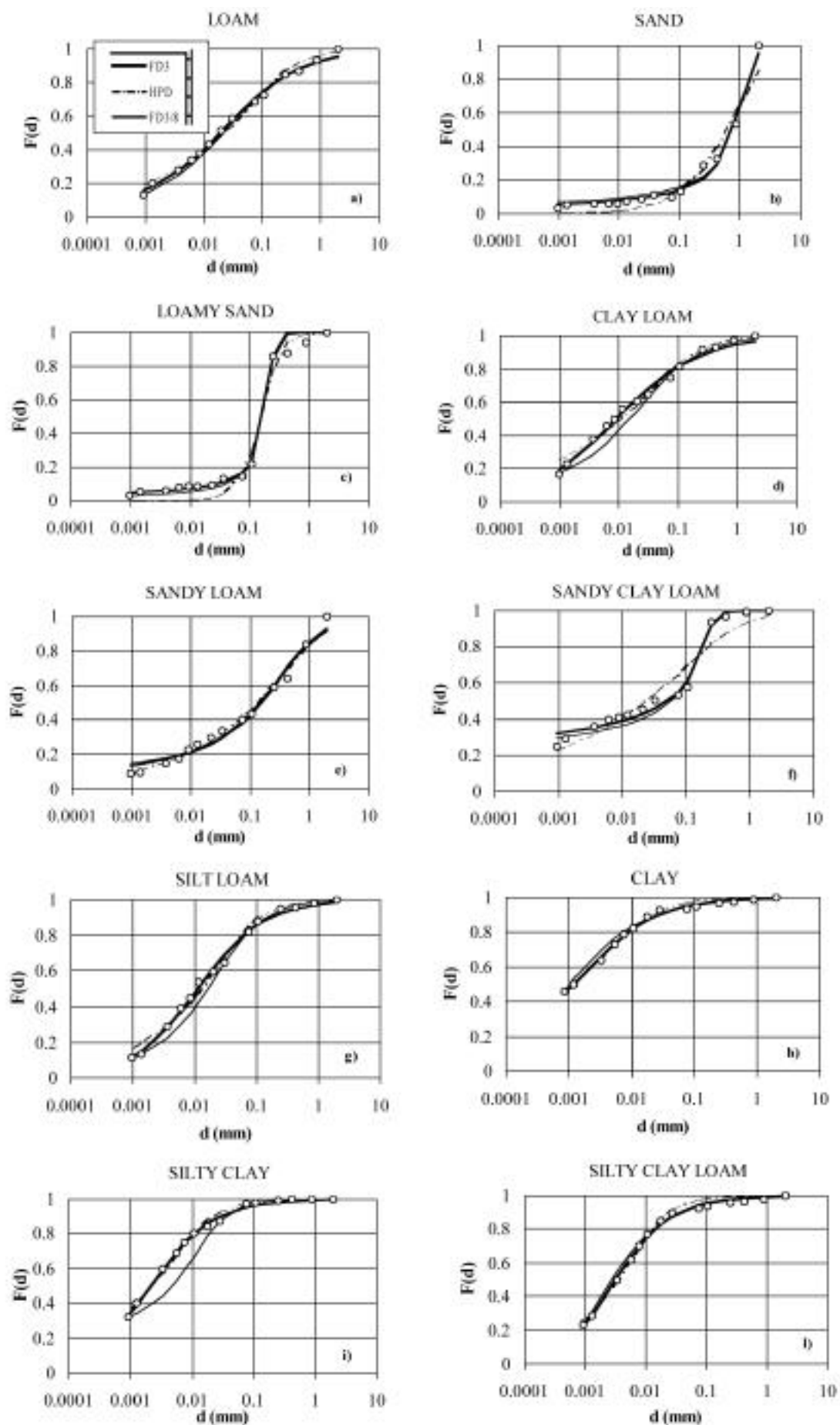


Fig. 5 - Comparison between the measured particle size distribution and the F (F3 and F3/8 scenarios) and HP models for some selected soil samples.

4. Conclusions

The aims of this study were to i) evaluate the ability of the HP and F models to fit 243 measured PSDs from a wide range of soil textures in Sicily, and ii) test the effect of the number of measured particle diameters on the fitting of the theoretical PSD. Each measured PSD was described by fourteen data points.

A three-parameter version of the F model, with fixed d_f (0.001 mm) and d_m (0.0001 mm) values, was found to be usable to describe the PSD.

Both the HP and F models performed better in fine-textured soils than in coarse-textured soils. However, the F model showed a better agreement with the measured PSDs than the HP model for a wide variety of soils. This result was not obvious, given that increasing the number of parameters cannot always guarantee improved performances. In conclusion, the F model with three estimated parameters was suggested for practical use in Sicily.

The parameter estimation of the F model was also carried out using only eight measured data pairs (six by sieving analysis and two by sedimentation analysis). Using a reduced experimental information increased the maximum discrepancy between measured and predicted frequencies and it also affected the ranking of model's performances by soil textural class and the detected goodness of the fitting in absolute terms. The number of measured particle-size fractions is a factor that should be taken into account when different investigations have to be compared.

Acknowledgements

All Authors analyzed the results and contributed to write the paper. The research was supported by grants of Progetto FEROS, Assessorato Agricoltura e Foreste, Regione Siciliana, and Progetto DIFA, Regione Sicilia. The authors wish to thank G.V. Di Piazza for the kind help in the initial stage of this investigation.

References

- Arya L.M., Paris J.F. A physic empirical model to predict the soil moisture characteristics from particle-size distribution and bulk density data. *Soil Sci. Soc. Am. J.*, 1981,45, 1023-1030.
- Buchan G.D. Applicability of the simple lognormal model to particle-size distribution in soils. *Soil Sci.*, 1989, 147,155-161.
- Buchan G.D., Grewal K.S., Robson A.B. Improved models of particle-size distribution: an illustration of model comparison techniques. *Soil Sci. Soc. Am. J.*, 1993, 57,901-908.
- Campbell G.S. 1985. *Soil physics with BASIC: Transport models for soil-plant systems*. Elsevier, Amsterdam.
- Fredlund D.G., Xing A. Equations for the soil-water characteristic curve. *Can. Geotech. J.*, 1994, 31,521-532.
- Fredlund, M.D., Fredlund D.G., Wilson G.W. An equation to represent grain-size distribution. *Can. Geotech. J.*, 2000, 37, 817-827.
- Fredlund M.D., Wilson G.W., Fredlund. D.G. Use of grain-size distribution for estimation of the soil-water characteristic curve. *Can. Geotech. J.*, 2002, 39,1103-1117.
- Gee G.W., Bauder J.W. Particle-size analysis. In: Klute, A. (Ed.), *Methods of soil analysis*. Part 1, 2nd ed. Madison, WI: ASA and SSSA, 1986, 383-411.
- Haverkamp R., Parlange J.Y. Predicting the water-retention curve from particle-size distribution: 1. Sandy soils without organic matter. *Soil Sci.*, 1986, 142,325-339.
- Hwang S.I. Effect of texture on the performance of soil particle-size distribution models. *Geoderma*, 2004, 123,363-371.
- Hwang S.I., Lee K.P., Lee D.S., Powers S.E. 2002. Models for estimating soil particle-size distribution. *Soil Sci. Soc. Am. J.*, 2002, 66,1143-1150.
- Jaky J. *Soil mechanics*. (in Hungarian) Egyetemi Nyomda, 1944, Budapest.
- Lassabatère L., Angulo-Jaramillo R., Soria Ugalde J.M., Cuenca R., Braud I., Haverkamp R. Beerkan estimation of soil transfer parameters through infiltration experiments – BEST. *Soil Sci. Soc. Am. J.*, 2006, 70:521-532.
- Minasny B., McBratney A.B. 2007. Estimating the water retention shape parameter from sand and clay content. *Soil Sci. Soc. Am. J.*, 2007, 71, 1105-1110.
- Nemes A., Wösten J.H.M., Lilly A., Voshaar J.H.O. Evaluation of different procedures to interpolate particle-size distributions to achieve compatibility within soil databases. *Geoderma*, 1999, 90,187-202.
- Saxton K.E., Rawls W.J., Romberger J.S., Papendick R.I. Estimating generalized soil-water characteristics from texture. *Soil Sci. Soc. Am. J.*, 1986, 50, 1031-1036.
- Skaggs T.H., Arya L.M., Shouse P.J., Mohanty B.P. Estimating particle-size distribution from limited soil texture data. *Soil Sci. Soc. Am. J.*, 2001, 65,1038-1044.
- Shiozawa S., Campbell G.S. On the calculation of mean particle diameter and standard deviation from sand, silt, and clay fractions. *Soil Sci.*, 1991, 152, 427-431.
- Shirazi M.A., Boersma L. A unifying quantitative analysis of soil texture. *Soil Sci. Soc. Am. J.*, 1984, 48, 142-147.
- van Genuchten M.T. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 1980, 44, 892-898.

SUMMARY

The soil particle-size distribution (PSD) is commonly used for soil classification and for estimating soil behavior. An accurate mathematical representation of the PSD is required to estimate soil hydraulic properties and to compare texture measurements from different classification systems. The objective of this study was to evaluate the ability of the Haverkamp and Parlange (HP) and Fredlund et al. (F) PSD models to fit 243 measured PSDs from a wide range of

soil textures in Sicily and to test the effect of the number of measured particle diameters on the fitting of the theoretical PSD. For each soil textural class, the best fitting performance, established using three statistical indices (*MXE*, *ME*, *RMSE*), was obtained for the F model with three fitting parameters. In particular, this model performed better in the fine-textured soils than the coarse-textured ones but a good performance (i.e., $RMSE < 0.03$) was detected for the majority of the investigated soil textural classes, i.e. clay, silty-clay, silty-clay-loam, silt-loam, clay-loam, loamy-sand, and loam classes. Decreasing the number of measured data pairs from 14 to eight determined a worse fitting of the theoretical distribution to the measured one. It was concluded that the F model with three fitting parameters has a wide applicability for Sicilian soils and that the comparison of different PSD investigations can be affected by the number of measured data pairs.

Keywords: Particle-size distribution, Particle-size distribution models, Soil physical properties.

List of symbols

- d particle diameter (mm);
 d_f parameter related to the amount of fines in a soil (mm);
 d_g scale parameter of the HP distribution (mm);
 d_m diameter of the minimum allowable particle size (mm);
 $F_i(d)$ frequency by weight;
 MXE statistical index (maximum error);
 ME statistical index (mean error);
 m shape parameter of the F model;
 N number of the measured pair of the PSD;
 n shape parameter of the HP distribution;
 n_F parameter related to the steepest slope on the curve for the F model;
 $P(d)$ fraction by weight of soil particles having a diameter less than or equal to d (mm);
 $P_i(d)$ probability corresponding to the frequency by weight $F_i(d)$;
 $RMSE$ statistical index (root mean square error);
 α parameter designating the inflection point of the F model.

